

ABSTRAK

Persoalan gender, khususnya *Lesbian, Gay, Bisexual, Transgender (LGBT)* menjadi masalah yang serius di Indonesia negara berideologi Pancasila. Opini masyarakat terhadap *LGBT* di Indonesia dapat diklasifikasikan menjadi tiga opini yaitu mendukung (pro), menolak (kontra), dan netral. Penelitian ini akan melihat pengaruh penggunaan algoritma *stemming* terhadap akurasi hasil pengklasifikasian menggunakan *Naïve Bayes*, dan *k-Nearest Neighbor*. Data yang digunakan adalah komentar – komentar berbahasa Indonesia yang di utarakan pada konten *YouTube* yang mengangkat topik *LGBT* di Indonesia. Melalui penelitian ini, harapannya mampu menjadi referensi penelitian terkait opini masyarakat terhadap *LGBT* di Indonesia melalui *YouTube*, sekaligus mengkaji kinerja klasifikasi dengan kebaruan menggunakan algoritma *Nazief Adriani* dalam proses *stemming*, untuk mengklasifikasikan opini masyarakat pada studi komparasi *Naïve Bayes*, dan *k-Nearest Neighbor classifier*, yang diharapkan hasil klasifikasi menjadi lebih akurat dan optimal.

Berdasarkan penelitian ini, *Stemming* menggunakan algoritma *Nazief Adriani* terbukti dapat meningkatkan akurasi klasifikasi hingga 8% menggunakan *K-Nearest Neighbor*, dan 2% menggunakan *Naïve Bayes*.

Akurasi terbaik klasifikasi komentar yang memuat opini terhadap *LGBT* di Indonesia menggunakan algoritma *K-Nearest Neighbour* adalah 53% dengan 80% data latih menggunakan *stemming*, dan nilai $K = 90-215$ Sedangkan akurasi terbaik klasifikasi menggunakan algoritma *Naïve Bayes* adalah 36% dengan 60% dan 70% data latih menggunakan *stemming*.

Kata kunci: *LGBT*, *Nazief Adriani*, *KNN*, *Naïve Bayes*

ABSTRACT

Gender issues, especially Lesbian, Gay, Bisexual, Transgender (LGBT) are a serious problem in Indonesia, a country with the Pancasila ideology. Public opinion on LGBT in Indonesia can be classified into three opinions, namely support (pro), reject (con), and neutral. This study will look at the effect of using the stemming algorithm on the accuracy of the classification results using Naïve Bayes, and k-Nearest Neighbor. The data used is Indonesian-language comments that are expressed on YouTube content that raises the topic of LGBT in Indonesia. Through this research, it is hoped that it can become a reference for research related to public opinion on LGBT in Indonesia through YouTube, as well as review the performance of classification with novelty using the Nazief Adriani algorithm in the stemming process, to classify public opinion in the Naïve Bayes comparative study, and the k-Nearest Neighbor classifier, It is hoped that the classification results will be more accurate and optimal.

Based on this research, Stemming using the Nazief Adriani algorithm is proven to increase classification accuracy up to 8% using K-Nearest Neighbor, and 2% using Naïve Bayes.

The best accuracy of the classification of comments containing opinions on LGBT in Indonesia using the K-Nearest Neighbor algorithm is 53% with 80% of the training data using stemming, and the value of $K = 90-215$ While the best accuracy of classification using the Naïve Bayes algorithm is 36% with 60% and 70% of training data using stemming.

Keywords: *LGBT, Nazief Adriani, KNN, Naïve Bayes*